“Agronomists don’t care about counterfactuals”

Evaluation research in Babati District

Ainsley Charles

Africa RISING Pre-planning workshop for ESA researchers and project teams
East Africa Hotel, Arusha, Tanzania, 6 August 2013

Agenda/Roadmap

- IE: what's the fuss? Why/why not? An overview (slides 3-10)
- Introducing improved hybrids to Babati farmers: evaluating WP2 (slides 11-16)
- Going forward/next steps/things to be determined (slide 17)

Questions of (potential) interest

- Can 'maize doctor' training mitigate disease and improve yields? (WP1)
- Do farmer field days increase uptake of new seed varieties? (WP2)
- Do these varieties (on their own) improve yields solely or when used in combination with local fertilizer? (WP2 and 4)
- Can farmer education challenge traditional culture/myths/legends (WP4)?
- Can changes in knowledge and behavior outcomes be sustained medium to long term
- Can these technologies raise consumption sufficiently to boost child and pregnant women outcomes?
- Livestock feed (WP3); Mycotoxin information (WP5)
- Should any of these experiments and trials be scaled up? (Long 2200, Matufa 7800, Sabilo 2900, Seloto 4200, ward hhs)

Landscape and AR cast of characters: Tanzania

- Real GDP growth 7.2 percent (2013/14, projected, IMF)
- Agricultural/total exports 17.1 percent
- Poverty incidence 33.6 percent (mainland, 2007), 40 percent (Zanzibar, 2009); 38.7 among agricultural households
- Agricultural growth (projected 6 percent) driven by increases in acreage (under cultivation), not productivity (per unit of land)
- Policy framework: ASDS '01, ASIP '05, *Kilimo Kwanza* policy
- Issues: underdeveloped infrastructure, insufficient extension, research
- Challenge: The search for *national* solutions BUT tested *locally* AND tested *comparatively* AND that are cost-effective

- Farming households
- Village authorities
- Extensionists
- Researchers/Implementers
- Agricultural scientists/Project manager
- Monitoring and evaluation team
- Donors, funders
- Governments

IE: Do I need it?

- Fear of program being shown not to work
- Fear of loss of contract, program
- Fear of being sidetracked with what's viewed as time-consuming and costly exercises
- 'I know my trial works – farmers (and monitoring data) say so!'
- Dislike of economists (!)

Why IE/what does it provide and to whom?

- Researchers, agencies, government officials, and policymakers
- A way to identify programs or program components that :
  - are working (and how well)
  - can be improved (mid-course corrections)
  - are ineffective
- Which investments are providing the most productive returns.
- A means to measure development effectiveness.
- Institutional development: evaluation culture, agencies, networks; programs survive political transitions.

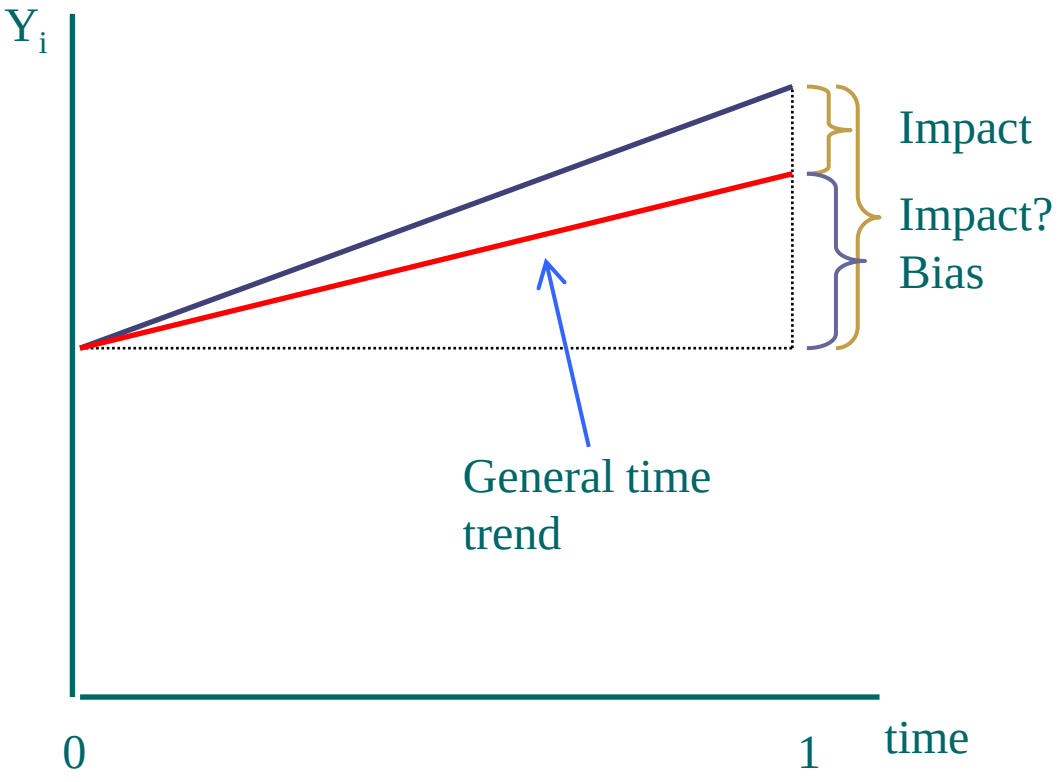
IE: definition, challenge, and measurement

- The ability to match the influence of a socioeconomic program to specific and anticipated outcomes, essentially to tell a story about its effectiveness.
- Social and economic programs are geared to improving the lives of beneficiaries
- Do not operate in isolation; many confounding factors
- Causally attribute changes in outcomes to the program while controlling for all other factors that may have played a role.

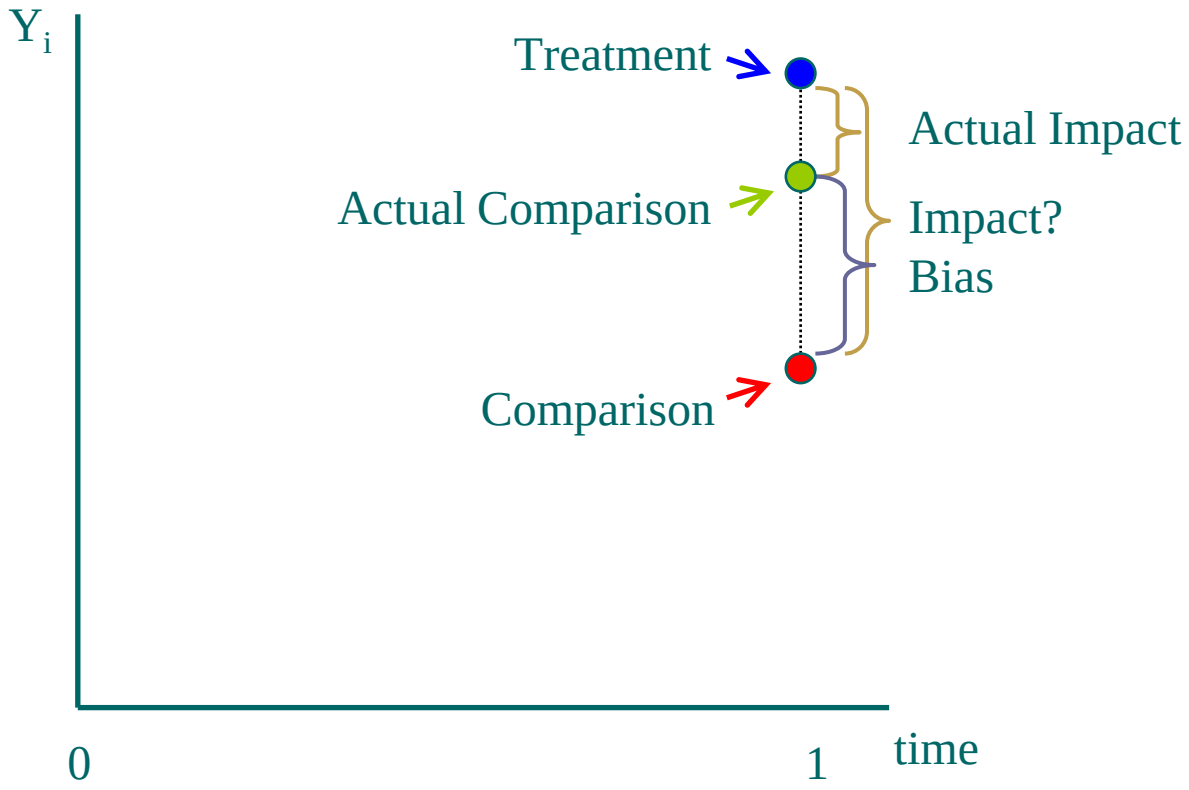
- The **impact** of a program is the difference between …
  - Outcomes achieved by program beneficiaries after having participated in the program and …
  - Outcomes **these same participants** would have achieved at the same moment in time if they had not participated in the program.
- The fundamental evaluation problem is that …
  - The second measurement, defined as the **counterfactual outcome**, is never observed. What can we do?
  - Alternative (and statistically rigorous) ways to measure counterfactual outcomes, substitute for the missing ones.

IE: Solving the missing data problem…maybe?

Using participants: before v after



Using non-participants: with v without



Sidebar: Why monitoring is necessary but not sufficient

- Monitoring …
  - Tracks indicators over time (program administrative information)
  - Collects data ONLY for participants
  - Permits descriptive analysis
  - Tells whether inputs are being implemented and outputs achieved as planned
- Impact evaluation …
  - Tracks average outcomes over time (surveys) and indicators (program information)
  - Collects data for participants and non-participants (although these may enter the program later on)
  - Permits causal/attribution analysis
  - Tells whether the program causally influences outcomes and why/by what mechanism

IE: Solving the same problem... just carefully

Experimental design/randomization

- Steps:
  1. Determine eligibility
  2. Collect baseline data
  3. Randomly assign eligible individuals to treatment and control group
  4. Treatment group receives program while control does not (at least in short run)
  5. Collect follow-up data
- Ensures average characteristics of both groups are the same
- Outcomes of the control group substitute for the missing counterfactual outcomes
- Difference in mean outcomes provide estimate of program impact

Non-experimental design

- If no experiment…what can be done?
- Need data with participants and non-participants,
- Latter useful for creating a counterfactual group with a similar range of characteristics as for participants (to mimic randomization)
- Collect baseline and follow-up data
- Depending on the data and the way the program is assigned, chose from a menu of empirical methods to calculate impact (again, as the difference in mean outcomes)

Part Two: IE in Babati District

- Potential research questions:
  - Are there barriers to adopting improved seed varieties for maize-based agriculture? If so, how might they be overcome?
    - Credit/capital; risk aversion; culture/tradition
  - Can the adoption of improved seeds on its own improve agricultural yields in Babati District? Or are complementary inputs, such as extension or fertilizer, required?
  - What is the impact of providing specific inputs on resource allocation across the agricultural value chain, on agricultural production, and on household welfare?
  - What's the most appropriate research design for testing/answering these questions?
    - Individual versus community-level randomization
    - Contamination/indirect treatment

WP2: Possible evaluation designs

- Intervention-1: Out of the 400 farmers at the July field days in Long, Sabilo, and Seloto villages, 200 are randomly assigned to receive vouchers for improved seeds (partial or full subsidy?) or Minjingu mazao (fertilizer), and 200 nothing
- Intervention-2: Among 800 farmers for next AR phase, randomly assign combinations of improved seeds, fertilizer, and production contracts among groups of 200; 200 receive the initial field day only
- Outcomes of interest: shock resilience, consumption, yield, nutrition and food security, health/anthropometrics, knowledge and behaviour

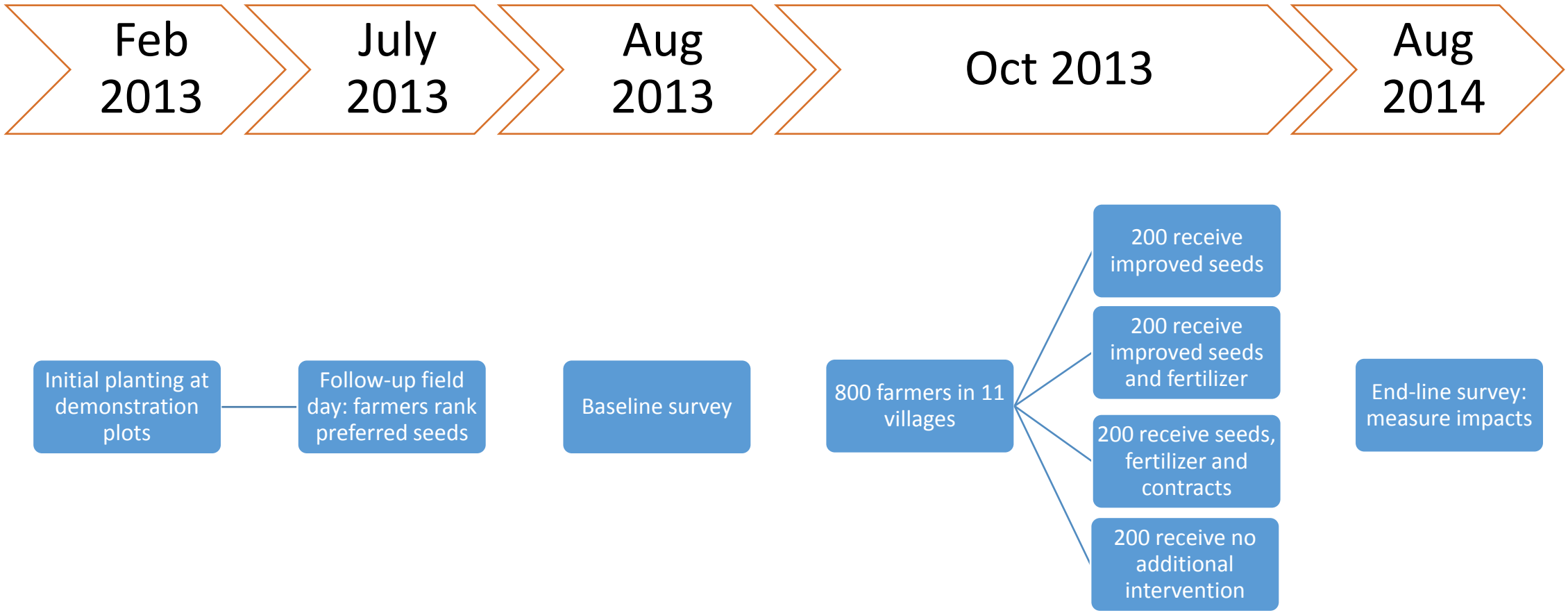
What might it look like?
A 'mini-IE' in operation: Long, Sabilo and Seloto

- Allocating inputs (tbd) to field day attendees by lottery
  - transparent, high buy-in
  - partial or full subsidy?
- Maximizing attendance: two days, village sensitization, local leader encouragement, researchers in attendance
  - 300-400 farmers
- Combined with follow-up to demonstrations (researchers), ranking of hybrids (farmers)
- Potential problems/challenges:
  - Too small sample
  - Contamination (now: individual-level randomization; later: work package interventions)

A ‘full’ IE of WP2: A collaborative approach

- A 1-2 day IE planning clinic/workshop: build IE partnerships
- Expanding the cast of characters: add an IE research team
- Scientific research team to pick from menu
- IE partners to pay for data collection
- Assumptions/Challenges:
  - Projects ready for full rollout October 2013
  - Sampling: able to draw a random sample from complete listing of farming households eligible to participate in AR

WP2: Timeline and design of an evaluation



Best results

- Build capacity.
- Promote learning in dissemination seminars and workshops.
- Contribute to the research literature with briefs and journal articles.
- Show the ministry of agriculture it works: Gain credibility for the program.
- Strengthen evidence-based culture for policy making.
- Have USAID expand the program.
- Have Government of Tanzania take up/over/scale up the program!

Next steps/things to be determined

- Other work packages: integrating IE
- Within each work package, determining specific technologies, trainings, delivery modalities to evaluate
- Results framework/FtF indicators: Monitoring inputs required
- Results framework/FtF indicators: How can IE help?

Maybe we all care about counterfactuals after all!

- The End -